

Preservação digital de longo prazo em instituições patrimoniais

Reutilização e adaptação de metadados

Fernanda Maria Guedes de Campos

Biblioteca Nacional
Campo Grande, 83
1749-081 Lisboa
Tel: 217982022
E-mail: fcampos@bn.pt

Maria de Lurdes Saramago

Biblioteca de Biologia
Faculdade de Ciências da Universidade de Lisboa
Campo Grande, Ed. C2, 2º Piso
1749-016, Lisboa
Tel: 217500000, Ext. 22220
E-mail: lurdes.saramago@fc.ul.pt

▪ RESUMO

O meio digital tem vindo a definir-se ao longo dos últimos anos, como preferencial durante a criação e troca de informação, contudo verifica-se ainda a existência de algumas limitações no que diz respeito à sua preservação para que possa ser recuperado dentro de 20 ou 50 anos.

Cada recurso digital armazenado deve ser enquadrado por metadados de forma a que a sua descrição para preservação futura esteja estruturada. Os metadados de preservação devem conter informação descritiva de carácter bibliográfico, estrutural e administrativo. Chamamos a atenção para o modelo OAIS (Open Archive Information System), que tem vindo nos últimos 4 anos a ser adaptado por um vasto grupo de comunidades, dadas as suas potencialidades para um enquadramento de aplicação alargada. Este modelo encontra-se em fase de auscultação de resultados e revisão.

Analisa-se os resultados do grupo de trabalho PREMIS (Preservation Metadata: Implementation Strategies Working Group) subsidiado pela OCLC/RLG e instalado na Biblioteca do Congresso. Este grupo tem vindo a observar e discutir a selecção e adaptação de matrizes de metadados consideradas eficazes dentro das comunidades.

A actividade do grupo ICABS-CDNL (IFLA-CDNL Alliance for Bibliographic Standards) também ele envolvido nesta matéria constitui de igual forma motivo de estudo e desenvolvimento nesta comunicação.

Pretende-se que os esquemas de metadados de preservação e seus conteúdos sejam exaustivos e percorram todo o ciclo de vida dos recursos digitais. Uma parte será criada através de software, sem recurso ao elemento humano, mas outra necessita de ser processada por pessoal especializado. Nenhum esquema de metadados é perfeito e deverá sempre ponderar-se a possibilidade de interoperabilidade através do estabelecimento de *crosswalks*. O esquema RDF permite a reutilização e permuta de diversos vocabulários, contudo é indispensável o

estabelecimento prévio de critérios para a sua reutilização.

▪ **PALAVRAS-CHAVE:** Preservação digital / Metadados / OAIS / PREMIS

▪ **INTRODUÇÃO: CONCEITOS, CONTEXTOS E ESTRATÉGIAS**

Ainda que os termos de preservação digital de longo prazo sejam usados já há, pelo menos, uma década, parece-nos importante clarificar, ainda hoje, o que significam: O termo preservação digital é usado em diferentes circunstâncias mas existem, essencialmente, dois contextos básicos para ele.

Por um lado, temos a preservação de informação contida num suporte através da digitalização da respectiva imagem. Por outro lado, temos a preservação digital enquanto armazenagem, manutenção e acesso ao recurso digital a longo prazo. Muita bibliografia (e também muitos projectos) tratam a preservação digital na primeira asserção que indicámos e há, ainda, relativamente pouca literatura e exemplos concretos que documentem a segunda (e mais abrangente) definição.

Quanto ao significado do “longo prazo”, trata-se de uma posição consciente, assumida praticamente desde o início da produção de recursos digitais e digitalizados, em tudo semelhante à posição que bibliotecas e arquivos, com funções patrimoniais, assumem no respeitante aos seus acervos: preservar para um futuro não de décadas mas de centúrias ou milénios.

No entanto, se a posição de princípio é idêntica, o mesmo não se poderá dizer dos métodos e práticas para levar a efeito o acto da preservação a longo prazo do património digital, conceito que, aliás, nesta introdução também vale a pena rever.

Património é definido nos documentos da UNESCO como o nosso legado do passado que possuímos hoje e que queremos passar às gerações futuras. O património é algo a que se atribui significado de geração em geração porque tem um determinado valor. O património digital é constituído por materiais produzidos no contexto informático de valor durável e

que se pretende passar às gerações vindouras. Nem todos os materiais digitais têm valor patrimonial mas para os que têm será necessário garantir a sua preservação pois só ela assegura a continuidade e a respectiva acessibilidade.

As ameaças mais significativas a esta continuidade do património digital têm a ver, essencialmente, com a perda dos meios de acesso e assim, a propósito da preservação dos recursos digitais é mantê-los acessíveis garantindo a integridade do seu conteúdo.

Que materiais devem ser preservados?

Este é sem dúvida, um aspecto que levanta, permanentemente múltiplas discussões. Os princípios de “importância” e de “efemeridade” que, de algum modo têm vindo a ser aplicados em bibliotecas e arquivos têm de ser reequacionados à luz dos novos tipos de documentos e formas de preparar e difundir informação. Por outro lado, há sempre que ter em conta as necessidades das comunidades de utilizadores que condicionam os critérios de selecção a aplicar, sendo que, no limite há sempre que definir princípios e responsabilidades institucionais pela preservação com a respectiva criação das condições de continuidade do acesso aos recursos.

De uma forma muito esquemática podemos apontar as principais linhas estratégicas que a preservação digital requer:

- Trabalhar com os produtores (criadores e/ou distribuidores) de modo a aplicar normas que garantam, à partida (isto é, no momento da criação) os meios de acesso, reduzindo assim o leque de problemas que a posteriori, são difíceis de resolver.
- Reconhecer que não é prático, viável e porque não dizê-lo ... desejável, preservar tudo, seleccionando mediante critérios estabelecidos, aquilo que deve ser preservado, mesmo com risco de perda de informação para os vindouros.
- Armazenar os recursos em mais do que um local seguro.
- Controlar os recursos utilizando metadados estruturados e outros meios, de forma a facilitar o acesso, garantindo, ao mesmo tempo, o apoio ao processo de preservação.
- Proteger a integridade e a identidade dos dados.
- Escolher normas apropriadas para providenciar o acesso, no ambiente tecnológico existente e emergente.
- Organizar os programas de preservação sempre em cooperação com outras entidades, nacionais ou internacionais, a fim de poupar esforços e recursos humanos e financeiros, gerir com boas práticas e alcançar os objectivos de preservação digital de maneira proactiva, holística e segura.

Do ponto de vista da organização e das boas práticas em preservação digital, estas orientações estratégicas têm

vindo a ser seguidas. Hoje já não se considera que as questões de preservação dizem só respeito a preservar os bits e a lidar com a obsolescência dos media e dos formatos dos ficheiros. Verificamos, aliás, que o reconhecimento da importância da preservação digital assenta no pressuposto de que é necessário desenvolver repositórios digitais de confiança que assegurem não só as migrações mas também a preservação da informação relevante sobre os objectos digitais por forma a preservarmos os seus “comportamentos” à medida que vão migrando, ou seja, a estrutura do objecto digital e o seu contexto, garantindo assim a sua acessibilidade no futuro.

Nesta perspectiva, tem sido absolutamente fundamental para a construção de repositórios digitais, o modelo de referência OAIS (Open Archive Information System), consignado em norma ISO. Com efeito, o modelo OAIS ainda que não seja um plano de implementação, providencia um enquadramento para a arquitectura e operacionalidade de um repositório digital, avultando no modelo a questão fundamental dos metadados. A preservação a longo prazo dos objectos digitais num repositório requer que se desenvolvam e mantenham metadados descritivos, estruturais e administrativos desses objectos.

Este aspecto que indicámos anteriormente será desenvolvido com mais pormenor nesta comunicação, onde serão apresentadas experiências práticas que acentuam a grande tendência que actualmente se verifica no domínio prático da preservação digital: a cooperação sobretudo no desenvolvimento de sistemas distributivos com responsabilidades partilhadas, em vez do modelo mais antigo do controlo centralizado ou das experiências isoladas.

A cooperação, por sua vez, significa hoje e cada vez mais a reutilização dos metadados, só possível pela aplicação de normas e formatos comuns. As normas são desde logo, a chave para criar um recurso que seja interoperável, ou seja, utilizável por outros sistemas, bem como sustentável ao longo dos tempos. Depender de software e formatos proprietários que não são compatíveis com normas comuns pode ditar a necessidade de migrar mais frequente e isoladamente. No respeitante aos metadados associados ao objecto digital, a visão estratégica deve ser a mesma: procurar, com a base estruturante de um modelo como o OAIS, disponibilizar os dados indispensáveis para conhecer, manipular e aceder aos recursos disponibilizando essa estrutura num princípio de partilha que é, afinal, a consequência directa – pelo menos para as instituições patrimoniais como as bibliotecas sobretudo nacionais, dos princípios de partilha de registos bibliográficos que nortearam a vida daqueles organismos desde os anos 70 do século XX e pelos quais foi indispensável criar normas de descrição e de formato que ainda hoje utilizamos na base dos trabalhos biblioteconómicos. A preservação digital não é, nem deve ser, uma linha de trabalho exógena à missão fundamental das instituições patrimoniais. O seu objectivo pode ter mais ou menos a ver com a acessibilidade a longo prazo e menos com a salvaguarda de um objecto enquanto forma física, como acontece com a preservação do património analógico.

Porém, não pode ser encarada como uma tarefa marginal, a desempenhar exclusivamente por agentes que não os curadores do património. Cada vez mais, as competências em preservação digital devem ser absorvidas pela instituição e seus agentes, desenvolvendo-se novas competências ao nível da produção de metadados e da gestão dos repositórios digitais.

▪ METADADOS DE PRESERVAÇÃO

Na generalidade pretende-se que os metadados contenham informação bibliográfica, estrutural e ainda outra administrativa acerca de um dado recurso. Quando falamos de metadados de preservação, estes resultam da combinação dos metadados descritivos já existentes aos quais são acrescentados elementos estruturais e administrativos que documentem p. ex., o sistema operativo, o formato dos ficheiros, a estrutura do conteúdo digital, um histórico de todas as acções de preservação exercidas sobre o recurso, histórico de custódia, informação sobre autenticidade e integridade, responsabilidades e direitos sobre os recursos, etc.

Os metadados administrativos e estruturais são considerados essenciais para a preservação de longo prazo.

A orientação dada pelos responsáveis dos repositórios em geral vai no sentido dos criadores de recursos digitais anexarem em conjunto com os seus trabalhos os metadados que permitam à partida documentar o recurso desde a sua criação. Na falta dessa informação é obrigação do repositório que os recebe suprir essa falha e gerar ele próprio esses metadados.

Todo o ciclo de vida dos recursos digitais é obrigatoriamente documentado e ainda são acrescentados metadados acerca dos metadados seleccionados.

É fundamental que o repositório defina e adopte um conjunto mínimo de metadados, estruturados de forma a acompanhar os recursos ao longo das suas sucessivas migrações.

É responsabilidade do repositório gerir e favorecer a utilização e reutilização dos metadados desde a sua ingestão de forma a assegurar os objectivos de preservação contemporâneos e futuros. Tanto no caso de recursos estáticos como dinâmicos, significa que pode vir a ter que existir um trabalho de actualização e enriquecimento dos metadados de forma a mantê-los acessíveis e legíveis depois de sucessivas migrações e mudanças de ambientes tecnológicos.

▪ ESTADO DA ARTE

Em 2001 o grupo OCLC/RLG (Online Computer Library Center/ Research Library Center) apresentou os resultados de um estudo sobre o estado da arte dos metadados de preservação em diversas comunidades [1]. As comunidades estudadas eram a CEDARS (CURL Exemplars in Digital Archives Project), a NLA (National Library of Australia), a NEDLIB (Networked

European Deposit Library) e ainda a comunidade Harvard.

As conclusões deste estudo apontaram para uma grande convergência de pontos de vista e boas possibilidades de atingir uma plataforma de normalização.

Como principal ponto de convergência encontramos a utilização do modelo de referência OAIS (Open Archive Information System)[2], atrás referido, criado no âmbito da NASA – CCSDS, posteriormente adoptado como norma ISO N. 14721:2002 e que manifesta as seguintes intenções:

- Determinar o âmbito do repositório em função da comunidade de utilizadores
- Negociar com os produtores de informação
- Controlar a informação por forma a atingir os objectivos da preservação de longo prazo
- Assegurar que a informação preservada é legível pela comunidade de utilizadores sem necessitar recorrer ao produtor de informação
- Manter procedimentos e políticas documentadas
- Tornar a informação preservada acessível à comunidade de utilizadores

O modelo OAIS apresenta uma descrição de alto nível e não transmite pressupostos acerca do tipo de recursos nem das especificações tecnológicas empregadas. A comunicação é estabelecida através de pacotes de informação em que um pacote de informação é um envelope conceptual onde se encontram encapsulados o recurso digital e os metadados de preservação. Os pacotes são 3: pacote para submissão, pacote para depósito e pacote para difusão.

Num contexto de preservação de metadados, a informação considerada mais relevante e indispensável encontra-se no pacote para depósito.

Este pacote é uma agregação de quatro blocos de objectos:

- 1) CI (Informação de conteúdo) - Informação acerca do conteúdo que o repositório tem a obrigação de preservar em conjunto com a informação de representação.
- 2) PDI (Informação descritiva para preservação) - Contém todos os elementos necessários para gerir a preservação da informação sobre o ambiente com que está associada. Este bloco divide-se em quatro tipos:
 - Informação acerca da referência – enumera e descreve os identificadores destinados à informação sobre o conteúdo de tal maneira que se tornem inequívocos, interna e externamente ao depósito (p. ex : ISBN, URN)
 - Informação acerca da proveniência - documenta a história da informação

sobre o conteúdo (p. ex. origem, histórico de custódia, acções e efeitos da preservação)

- Informação acerca do contexto - documenta as relações entre a informação sobre o conteúdo e o seu ambiente (p. ex. razões pelas quais foi criado, relações com outras informações de conteúdo, etc.)
 - Informação acerca da reparabilidade: documenta mecanismos de reparabilidade e autenticação usados para assegurar que o conteúdo da informação não foi alterado de forma não documentada (p. ex. checksums ou assinaturas digitais)
- 3) PI (Informação para empacotamento) - Envolve o objecto digital e os metadados associados numa unidade ou pacote.
 - 4) DI Informação descritiva - Facilita o acesso à informação sobre o conteúdo através das ferramentas de pesquisa e recuperação. A informação descritiva serve de input das ajudas à localização de repositórios e deriva tipicamente da informação sobre o conteúdo ou da informação descritiva para preservação.

Passados 4 anos encontramos diversas comunidades que continuam a aplicar o modelo OAIS e vamos dar especial atenção ao grupo de trabalho da OCLC/RLG com o projecto PREMIS.

O projecto PREMIS

Em 2003 a OCLC e o RLG criam o grupo de trabalho internacional PREMIS (PREservation Metadata – Implementation Strategies). Este grupo inclui participantes dos Estados Unidos da América, Austrália, Gran-Bretanha, Alemanha, Holanda e Nova Zelândia. Os objectivos apresentados foram:

- Definir um conjunto de metadados chave, de aplicação alargada na comunidade de preservação digital e redigir uma lista explicativa de apoio a cada um desses metadados. A esse documento de trabalho foi dado o título de Data Dictionary [3];
- Avaliar estratégias alternativas de codificação, armazenamento e gestão dos metadados de preservação e também favorecer a intercomunicabilidade entre sistemas;
- Promover o aparecimento de programas piloto cujo objectivo seria o de testar as boas práticas do PREMIS;

A lógica do modelo PREMIS

Seguindo Priscilla Caplan & Rebecca Guenther [4] e também o relatório final do projecto, Data Dictionary [3], um recurso digital pode conter em si mesmo várias

representações, p. ex. um sítio Web pode conter uma página Web que pode incluir textos, gráficos, fotos, som, animação, etc.

O modelo PREMIS organiza-se em 5 tipos de entidades:

Objectos
Entidades intelectuais
Eventos
Direitos
Agentes

Em que um objecto é uma unidade de informação discreta em formato digital e que por sua vez pode ser englobada em 3 subtipos :

Ficheiros

Um ficheiro é considerado como uma sequência de bytes reconhecido por um sistema operativo. Um ficheiro pode ser composto por zero ou mais bytes e tem um formato, tem permissões de acesso e também incorpora alguns dados tais como o tamanho e a data da última modificação.

Fluxo de bites

No PREMIS, um fluxo de bits é considerado como um conjunto de dados num dado ficheiro, que pode ser contíguo ou não contíguo e que manifesta propriedades comuns para fins de preservação. Um fluxo de bites não tem autonomia se não estiver contido numa estrutura.

Representação

A representação é considerada como um conjunto de ficheiros, incluindo os metadados estruturais, necessários para restabelecer a natureza e o aspecto de uma entidade intelectual.

Uma entidade intelectual é um conjunto coerente descrito como uma unidade, p. ex: um livro, um mapa ou uma base de dados.

Eventos são entidades que agregam metadados acerca de qualquer acção ocorrida durante o ciclo de vida do recurso antes ou depois deste ser aceite e incluído no repositório.

Agentes são pessoas, organizações ou aplicações informáticas associadas às acções de preservação de um dado objecto digital

Os direitos são as permissões legais de cada objecto ou agente.

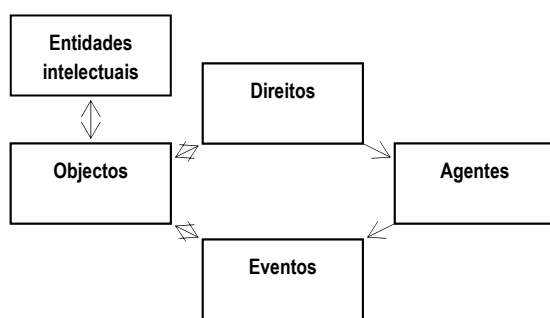
No que diz respeito à implementação, os parceiros PREMIS são bastante condicionados pelo modelo ao uso de padrões existentes e o espaço de desvio é muito reduzido, p. ex: metadados de uso local podem ser usados para alargar o grupo disponível mas os metadados já existentes não podem ser modificados.

O conjunto de campos de preenchimento obrigatório está relacionado com a longevidade dos recursos, que se pretende longa, e também com as transferências de um repositório para outro.

Dá-se como exemplo a norma Z39.87 (Technical Metadata for Digital Still Images) que reviu alguns dos seus elementos para poder harmonizar-se com o esquema PREMIS.

Em Portugal o projecto RODA do IANTT referido acima experimenta alinhar o modelo PREMIS com os esquemas de metadados EAD e METS. A organização do conjunto de esquemas de metadados baseia-se no PREMIS em que o esquema EAD é usado para a descrição da Entidade Intelectual. A norma Z39.87 é aplicada à descrição do objecto e o esquema METS estrutura os ficheiros e cria uma representação da entidade intelectual [5].

Seguindo o PREMIS, o armazenamento dos metadados num sistema de bases de dados tem a vantagem do acesso fácil, actualização rápida e possibilidade da criação de interrogações e relatórios em qualquer momento. Por outro lado, se os metadados se encontrarem em conjunto com os objectos também existem algumas vantagens pois as estratégias de preservação são aplicadas em simultâneo. A recomendação vai no sentido da redundância para o caso de metadados considerados de importância extrema.



Modelo PREMIS

A ingestão dos metadados no repositório

A maior parte dos repositórios, recebe grandes quantidades de recursos, desta forma a recomendação PREMIS vai no sentido da criação e uso do maior número possível de metadados gerados automaticamente. Para facilitar o processamento automático é recomendado o uso de vocabulários controlados pois a utilização descontrolada de diferentes vocabulários impede a interoperabilidade. A fonte de cada vocabulário deve ser anotada quando se importam metadados.

AS ACTIVIDADES DO GRUPO ICABS/CDNL

Em 2004-2005 a Biblioteca Nacional da Holanda (Koninklijke Bibliotheek) desenvolveu uma investigação para a IFLA-CDNL Alliance for Bibliographic Standards (ICABS) no sentido de sistematizar o conhecimento dos desenvolvimentos de padronização, a nível internacional, no que diz respeito à preservação digital [6].

Esta pesquisa envolveu as bibliotecas nacionais da Austrália, Austria, Canadá, China, Dinamarca, França, Alemanha, Japão, Holanda, Nova Zelândia, Portugal, Suécia, Suíça, Reino Unido e Estados Unidos da América.

Todas estas bibliotecas construíram ou têm intenção de construir o seu repositório digital baseadas no modelo OAIS, dada a capacidade do modelo ser adaptável a qualquer comunidade. Esta opção demonstra também a boa vontade deste conjunto de grandes bibliotecas nacionais de querer seguir uma norma que facilite a comunicação entre as comunidades.

Dado que um dos problemas com que lidam os repositórios é a ampla variedade de formatos, também foi averiguado no âmbito deste estudo quais os formatos mais comuns e que são : TIFF para imagem, PDF e XML para texto, HTML para sítios Web, WAV para áudio. Também se encontram referidos outros formatos, mas de uso menos frequente e que são: ZIP, PNG, ISO, JPEG, GIF, TeX, PS, EPS, AI, RTF, TXT e ASCII.

No que diz respeito aos metadados de preservação, estes são em grande parte gerados automaticamente, mas existe também algum trabalho manual.

As bibliotecas armazenam os metadados seguindo 3 metodologias diferentes. A de uso mais corrente é a manutenção em bases de dados em linguagem XML, separadas dos recursos, adoptada pelas bibliotecas da Holanda, Austrália, Canadá, Alemanha e Nova Zelândia. Neste caso as ligações são asseguradas por identificadores persistentes tais como NBN ou URN. Um segundo método é o de separar os metadados mas mantê-los no mesmo repositório, este é seguido pela Áustria e pela Dinamarca. Por sua vez a BNF mantém os metadados em conjunto com os ficheiros.

▪ SISTEMAS DE METADADOS MAIS UTILIZADOS

Todas estas bibliotecas utilizam mais do que 1 esquema de metadados e o METS é usado por todas.

O sistema METS

O sistema de metadados METS é uma norma para codificação de metadados descritivos, administrativos e estruturais de recursos digitais que utiliza a linguagem XML. A norma é mantida pelo Network Development and MARC Standards Office da Biblioteca do Congresso e tem sido desenvolvida como uma iniciativa da DLF (Digital Library Federation) [7] .

Esta iniciativa pretende construir um formato XML para codificação de metadados necessários para a gestão de bibliotecas digitais no âmbito de repositórios e também para permitir a permuta de recursos entre repositórios ou entre repositórios e utilizadores.

Dependendo da sua utilização, um documento METS pode ser usado como SIP (Submission Information Package), como AIP (Archival Information Package) ou mesmo como DIP (Dissemination Information Package) no âmbito do modelo de referência OAIS atrás referido.

Um documento METS é constituído por cinco secções :

- 1) Metadados descritivos – Esta secção pode consistir em apontadores para esquemas externos de metadados tais como em formato MARC, através de um OPAC ou ajudas EAD através de um servidor WWW.
- 2) Metadados administrativos – Fornecem informação acerca da fonte dos dados e acerca da proveniência, da história do recursos (migrações, transformações, etc.)
- 3) Grupos de ficheiros – Lista dos ficheiros que compõem o recurso em todas as versões.
- 4) Mapa de estruturas – Estrutura hierárquica dos recursos digitais e das ligações entre o conteúdo dos ficheiros e os metadados
- 5) Comportamento – Secção relativa a comportamento que pode ser usada para associar comportamentos executáveis com o conteúdo de um recurso cujos metadados pertencem ao esquema METS. Esta secção contém também um mecanismo modular de um código executável que implementa e instala comportamentos definidos abstractamente pelo interface.

Além do METS, também algumas bibliotecas recorrem ao uso do PREMIS Data Dictionary também já referido atrás, são elas a da Austrália, da Áustria, da Nova Zelândia e dos Estados Unidos da América.

Não é possível falar em metadados para objectos digitais quer sejam de carácter geral ou de preservação e omitir o esquema utilizado em âmbito mais alargado que é o Dublin Core.

Dublin Core

O Dublin Core é o formato de metadados para recursos digitais passível de aplicação alargada a qualquer comunidade temática. Foi projectado tendo como objectivo a recuperação dos recursos em HTML. Contém uma estrutura simples e flexível aplicável a recursos complexos. Tem capacidade para ser representado através de variadas sintaxes e podemos encontrar uma codificação para os elementos do Dublin Core em XML usando RDF [8].

Estabelecido no âmbito da OCLC/NCSA, é composto por um conjunto de 15 elementos padrão, permitindo a inclusão de elementos adicionais para atender às particularidades de cada comunidade. Foi estabelecido por consenso de um grupo internacional e interdisciplinar de profissionais – bibliotecários, analistas, linguistas, museólogos, entre outros, e é utilizado para descrever uma variedade de recursos existentes na Internet. Pode ser considerado como um meio de comunicação e de busca de informação para a Internet. O DC tem sido adoptado por importantes instituições e também como padrão nacional em agências governamentais americanas. Apresentamos em seguida o conjunto de elementos definidos para aplicação pelo Dublin Core.

- Título - título dado ao recurso
- Criador - entidade principal responsável pela elaboração do conteúdo do recurso
- Assunto - assunto referente ao conteúdo do recurso
- Descrição - descrição do conteúdo do recurso
- Editor – entidade responsável pela difusão do recurso
- Outros responsáveis - entidade responsável por contribuições ao conteúdo do registo
- Data - data da criação ou de actividade de preservação no ciclo de vida do recurso
- Tipo - natureza ou género do conteúdo do recurso
- Formato - manifestação física ou digital do recurso
- Identificador - identificação única não ambígua do recurso
- Fonte - referência a um outro recurso do qual deriva o presente
- Língua - língua do conteúdo intelectual do recurso
- Relação - referência para outro recurso do qual o presente recurso é derivado ou está associado
- Cobertura - âmbito espaço-temporal do conteúdo do recurso
- Direitos - Informação sobre os direitos legais do recurso e seu uso

No âmbito da preservação digital o esquema DC é frequentemente criticado por ser muito simples, mas os seus 15 elementos podem ser expansíveis a mais de 50 e num contexto web é sem dúvida um expoente de interoperabilidade. Podemos contudo condicionar que este esquema é mais eficaz no âmbito descritivo.

CONCLUSÃO

Um repositório na Internet é inevitavelmente um novo tipo de organização, com competências para responder aos problemas do próprio ambiente. Este pode não estar sediado num espaço físico, pode existir distribuído entre instituições em muitas localidades geográficas através de uma rede global.

A preservação digital é um problema com características de urgência na sociedade moderna e ao tentar assegurar a longevidade de uma herança cultural digital o homem enfrenta os desafios do novo paradigma relacionados com a efemeridade dos recursos e das tecnologias.

Focalisámos este trabalho sobre o estado da arte das experiências e implementações internacionais de metadados de preservação digital.

A incidência sobre os metadados administrativos e estruturais é relevante, pois é neste espaço que vamos encontrar as descrições dos métodos e das estratégias tomadas para preservação. Os metadados descritivos destinam-se fundamentalmente às fases de acesso e

estão para os recursos digitais como os formatos MARC (Machine Readable Cataloging) para os recursos bibliográficos tradicionais.

Pudemos constatar também, que o DC tem excelente capacidade descritiva para servir o acesso, mas já o mesmo não se passa no que diz respeito à preservação e ao cumprimento da função dos metadados administrativos e estruturais de documentar a história do recurso com estratégias de preservação, componentes de hardware e software, etc. etc.

Ao atribuir-lhe qualificadores em larga escala é possível tornar o conjunto dos elementos do DC eficiente em relação ao que cada comunidade pretende, contudo acabamos por transformar o que é simples em algo muito complexo.

Perante a existência dos 2 muito interessantes e eficazes sistemas de metadados que mencionámos neste trabalho, o Data Dictionary do PREMIS e o METS pensamos que a capacidade de reutilização dos seus elementos viabiliza a sua adaptação a qualquer comunidade.

Tendo em atenção aquilo que para trás ficou dito, podemos confirmar a urgência e a necessidade de cada vez mais as bibliotecas patrimoniais e não só, se organizarem em consórcios e desta forma partilharem os recursos necessários à criação e a uma boa gestão dos repositórios digitais. No caso particular dos metadados de preservação o caminho a seguir vai no sentido da sua reutilização, da sua readaptação e sempre no seguimento das normas internacionais existentes com vista a facilitar a intercomunicabilidade entre sistemas.

REFERÊNCIAS

1. OCLC/RLG – Preservation metadata for digital objects: a review of the state of the art: a white paper. 2001. [Em linha] [Consult. 13.Dec.2006]. Disponível em: http://www.oclc.org/research/projects/pmwg/presmeta_wp.pdf
2. CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS– *Reference Model for an Open archive Information System* (OAIS), Blue Book (CCSDS 650.0-B-1), 2002. [Em linha] [Consult. 07.Dec. 2006]. Disponível em: <http://nost.gsfc.nasa.gov/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>
3. OCLC/RLG – Data dictionary for preservation metadata: final report of the PREMIS Working Group. Dublin, Ohio, 2005 [Em linha] [Consult. 07.Dec. 2006]. Disponível em: <http://www.oclc.org/research/projects/pmwg/premis-final.pdf>>
4. CAPLAN, P., GUENTHER, R. – Practical preservation: the PREMIS experience. *Library Trends*.54:1 (2005), 111-124. [Em linha] [Consult. 13.Jan. 2007]. Disponível em: http://www.loc.gov/standards/premis/caplan_guenther-librarytrends.pdf
5. IAN/TT - RODA - Repositório de objectos digitais autênticos . [Em linha] [Consult. 07.Dec. 2006]. Disponível em: http://roda.iantt.pt/?q=pt/sobre_nos
6. VERHEUL, Ingeborg – *Networking for digital preservation* : current practice in 15 National Libraries. München, K. G. Sauer, 2006. ISBN 3-598-21147-8 (IFLA Publications ; 119)
7. Metadata Encoding & Transmission Standard [em linha] [Consult. 19.Jan.2007]. Disponível em : <http://www.loc.gov/standards/mets/>
8. BECKETT, Dave, MILLER, Eric, BRICKLEY, Dan - Expressing simple Dublin Core in RDF/XML. Dublin Core Metadata initiative [em linha][Consult. 19.Jan.2007]. Disponível em : <http://dublincore.org/documents/2002/07/31/dcmes-xml/>